ISSN: 2339-2541
JURNAL GAUSSIAN, Volume 5, Nomor 3, Tahun 2016, Halaman 373-382
Online di: http://ejournal-s1.undip.ac.id/index.php/gaussian

ANALISIS PENGARUH JUMLAH UANG BEREDAR DAN NILAI TUKAR RUPIAH TERHADAP INDEKS HARGA SAHAM GABUNGAN MENGGUNAKAN PEMODELAN REGRESI SEMIPARAMETRIK KERNEL

Deden Aditya Nanda[1], Suparti[2], Abdul Hoyyi[3]
[1]Mahasiswa Jurusan Statistika FSM Universitas Diponegoro
[2,3]Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

ABSTRACT

Stocks are one of the many forms of investment chosen by the investor. Investors can use Composite Stock Price Index (CSPI) as one of the indicators that show the movement of stock prices. CSPI fluctuates every day, where one of the causes are macroeconomic factors. Therefore needs to be done a proper analysis to model the CSPI and the factors that influence it. This study is using 1 parametric component variable (money supply) and 1 nonparametric component variable (exchange rate the rupiah against the dollar). So that proper modeling is semiparametric regression. Nonparametric component will be using kernel regression method by selecting the optimal bandwidth using a generalized cross validation method (GCV). This study uses monthly data. Data *in sample* is used as much as 68 data that is taken from Januari 2010 to August 2015, meanwhile *out sample* that is used as much as 6 data from September 2015 to February 2016. Based on the results of the analysis that has been done, the best kernel semiparametric regression model is using gaussian kernel function with bandwidth is around 47.94 and GCV=34675.27047. Determination coefficient value is 0.9781. Evaluation result of the model for value of Mean Absolute Percentage Error (MAPE) data out sample is around 4,036%, which indicates that the model is very accurate.

Keywords: Composite Stock Price Index (CSPI), Semiparametric regression, Kernel, GCV

1. PENDAHULUAN

Perekonomian sebuah negara tidak lepas dari peranan penting pasar modal. Salah satu fungsi yang dijalankan pasar modal adalah fungsi ekonomi, yaitu fungsi yang menyediakan sarana untuk mempertemukan dua kepentingan, yaitu pihak yang memiliki kelebihan dana (*investor*) dan pihak yang memerlukan dana (*issuer*). Dengan kata lain, pasar modal merupakan sarana pembentuk modal dan akumulasi dana jangka panjang untuk meningkatkan partisipasi masyarakat dalam penggerakan dana guna menunjang pembiayaan pembangunan nasional. Fungsi lainnya dalam pasar modal adalah sebagai sarana berinvestasi dalam berbagai bentuk, salah satunya dalam bentuk saham[7]. Dalam melakukan investasi investor membutuhkan informasi mengenai perkembangan saham yang akan menentukan bagaimana risiko dan imbal balik ke depannya. Informasi tersebut salah satunya dapat berupa Indeks Harga Saham Gabungan[8]. Indeks Harga Saham Gabungan (IHSG) adalah suatu nilai yang digunakan untuk mengukur kinerja gabungan seluruh saham yang tercatat di bursa efek[13].

Harga saham berfluktuasi setiap hari dan para pengamat meyakini bahwa fluktuasi harga saham ini sangat dipengaruhi oleh berbagai faktor di luar pasar modal dimana salah satunya adalah faktor makroekonomi. Faktor-faktor makroekonomi yang dimaksud antara lain jumlah uang beredar, kurs rupiah terhadap dollar, inflasi, tingkat suku bunga, dan harga minyak mentah dunia[11]. Berdasarkan studi awal penelitian, peneliti membuat *scatterplot* antara IHSG beserta faktor-faktor yang mempengaruhinya. Dilihat dari *scatterplot* setiap faktor yang mempengaruhi IHSG, terdapat faktor yang memiliki pola tertentu (komponen parametrik) yaitu jumlah uang beredar dan terdapat pula faktor yang tidak memiliki pola tertentu (komponen nonparametrik) yaitu tingkat inflasi, nilai tukar rupiah terhadap dollar, tingkat suku bunga, dan harga minyak mentah dunia. Dari beberapa faktor tersebut, peneliti tertarik untuk membahas Indeks Harga Saham Gabungan dengan dua faktor yaitu jumlah uang beredar dan kurs rupiah terhadap dollar. Karena menggunakan dua faktor dengan pola tertentu (linier) dan tak tentu, untuk mendapatkan model terbaik digunakan metode regresi semiparametrik. Pada penelitian ini, peneliti menerapkan fungsi kernel gaussian dan segitiga. Masalah dalam penelitian ini dibatasi pada jumlah data yang digunakan yaitu sebanyak 68 data dari bulan Januari 2010 sampai Agustus 2015. Tujuan dalam penelitian ini adalah menentukan pemodelan untuk IHSG dengan menggunakan regresi semiparametrik kernel.

.

2. TINJAUAN PUSTAKA

2.1 Indeks Harga Saham Gabungan

Indeks Harga Saham Gabungan atau IHSG (*Composite Stock Price Index*) adalah suatu nilai yang digunakan untuk mengukur kinerja gabungan seluruh saham yang tercatat di bursa efek. Maksud dari gabungan seluruh saham ini adalah kinerja saham yang dimasukkan dalam perhitungan seluruh saham yang tercatat di bursa tersebut[13]. Terdapat dua metode dalam perhitungan IHSG, yaitu metode rata-rata dan timbangan[1].

2.2 Faktor-Faktor yang Mempengaruhi Indeks Harga Saham Gabungan

IHSG dipengaruhi oleh banyak faktor, antara lain jumlah uang beredar dan kurs rupiah terhadap dollar[11].Pengertian uang beredar telah dibedakan pula menjadi dua pengertian, yaitu pengertian yang terbatas dan pengertian yang luas. Dalam pengertian yang terbatas uang beredar adalah mata uang dalam peredaran ditambah dengan uang giral yang dimiliki perseorangan, perusahaan, dan badan pemerintah. Dalam pengertian yang luas, uang beredar meliputi: (i) mata uang dalam peredaran, (ii) uang giral, dan (iii) uang kuasi[12]. Ketika jumlah uang yang beredar di masyarakat mengalami peningkatan maka orang akan cendrung berinvestasi pada saham, karena dengan itu mereka akan mendapatkan keuntungan sesuai ekspektasi. Sedangkan jika mereka berinvestasi pada bank, maka keuntungannya tidak sesuai dengan yang diinginkan karena pada saat itu suku bunga akan mengalami penurunan[11]. Sedangkan kurs adalah jumlah uang domestik yang dibutuhkan, yaitu banyaknya rupiah yang dibutuhkan, untuk memperoleh satu unit mata uang asing[12]. Perubahan kurs rupiah terhadap dollar memiliki pengaruh yang berbeda-beda untuk para emiten. Kenaikan kurs akan berdampak negatif terhadap emiten yang memiliki utang dalam dollar dan berorientasi lokal, tetapi akan berdampak positif terhadap emiten yang tidak memiliki utang dalam dollar dan berorientasi ekspor[11].

2.3 Regresi Parametrik

Analisis regresi parametrik mengasumsikan bahwa bentuk fungsi atau kurva regresinya sudah diketahui. Apabila pada *scatterplot* terlihat bahwa data memiliki kecenderungan mengikuti pola linier maka digunakan model parametrik linier. Berikut model regresi linier berganda jika diberikan n observasi dengan p variabel prediktor[5]:

$$y_i = \beta_0 + \sum_{k=1}^{p} \beta_k x_{ik} + \varepsilon_i \quad ; i = 1,2,\dots,\mathrm{n}\,; \ \mathrm{k} = 1,2,\dots,\mathrm{p} \qquad (1)$$

Dengan: y_i : Variabel respon pengamatan ke-i

x_{ik} : Variabel bebas ke-k pengamatan ke-i

β_k : Koefisien regresi pada x_k

β_0 : Intersep

ε_i : Residual ke-i

Jika ditulis dalam bentuk matriks, persamaan (1) dapat ditulis sebagai berikut:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \qquad (2)$$

2.4 Estimasi Model Regresi Parametrik

Metode yang digunakan untuk mengestimasi parameter model regresi adalah dengan meminimumkan jumlah kuadrat error atau yang sering dikenal dengan *Ordinary Least Square* (OLS) [4]. Berdasarkan persamaan (2), dengan metode kuadrat terkecil diperoleh estimasi β, yaitu sebagai berikut :

$$\sum_{i=1}^{n} \boldsymbol{\varepsilon_i}^2 = \boldsymbol{\varepsilon^T\varepsilon} = (\boldsymbol{Y}-\boldsymbol{X\beta})^{\boldsymbol{T}}(\boldsymbol{Y}-\boldsymbol{X\beta}) = \boldsymbol{Y^TY} - \boldsymbol{2\beta^TX^TY} + \boldsymbol{\beta^TX^TX\beta}$$

Dengan syarat $\frac{\partial\varepsilon^T\varepsilon}{\partial\beta} = 0$, maka diperoleh $\boldsymbol{-2X^TY + 2X^TX\hat{\beta} = 0}$

$$\boldsymbol{X^TY = X^TX\hat{\beta}}$$

$\frac{\partial^2\varepsilon^T\varepsilon}{\partial\beta\partial\beta^T} = \boldsymbol{2X^TX},$ merupakan matriks semidefinit positif. Jika $\boldsymbol{2X^TX}$ adalah matriks definit positif, maka $\hat{\beta}$ merupakan estimator kuadrat terkecil untuk β dengan:

$$\hat{\beta} = (\boldsymbol{X^TX})^{-\boldsymbol{1}}\,\boldsymbol{X^TY}$$

2.5 Regresi Non Parametrik

Konsep regresi nonparametrik berbeda dengan regresi parametrik. Pendekatan regresi nonparametrik merupakan suatu metode untuk menganalisis fungsi atau kurva regresi yang tidak diketahui[5]. Model umum regresi nonparametrik dapat dituliskan sebagai berikut:

$$y_i = m\,(t_i) + \varepsilon_i\,;\, i = 1,2,\ldots,n \qquad (3)$$

Dengan y_i : Variabel respon pengamatan ke-i
t_i : Variabel prediktor pengamatan komponen nonparametrik ke-i
m : Fungsi regresi yang tidak diketahui
ε_i : Residual ke-i

2.6 Fungsi Densitas Kernel

Mengetahui bentuk dari distribusi suatu data merupakan hal yang berguna dalam mengeksplorasi data. Histogram merupakan salah satu teknik yang biasa digunakan untuk mengetahui bentuk distribusi suatu data, namun tidak memberikan estimasi densitas yang mulus (*smooth*). Sehingga, dibutuhkan estimator penghalus dari fungsi densitas yang sebenarnya agar fungsi tersebut mulus. Salah satu estimator penghalus yang dapat digunakan yaitu estimator kernel yang merupakan pengembangan dari histogram[10]. Pendekatan kernel tergantung pada *bandwidth h*, yang berfungsi untuk mengontrol *smoothness* dari kurva estimasi. Pemilihan *Bandwidth* optimal menggunakan metode *Generalized Cross Validation* (GCV).

Fungsi kernel mempunyai sifat- sifat yaitu[2]:

1. $K(x) \geq 0$
2. $\int_{-\infty}^{\infty} K(x)\,dx = 1$
3. $K(-x) = K(x)$
4. $\int_{-\infty}^{\infty} x\,K(x)\,dx = 0$
5. $\int_{-\infty}^{\infty} x^2\,K(x)\,dx < \infty$

Jika diberikan data pengamatan $x_1, x_2, \ldots, x_n$ dari suatu distribusi dengan densitas f (tidak diketahui), maka estimator densitas kernel dengan fungsi kernel K dan h merupakan lebar kelas yang disebut *bandwidth* didefinisikan sebagai:

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad (4)$$

Ada beberapa fungsi kernel yang biasa digunakan, diantaranya adalah:

1. Kernel Gaussian

$K(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, -\infty < x < \infty$

2. Kernel Segitiga

$K\,(x) = \begin{cases} 1 - |x|, & |x| \leq 1 \\ 0\,, & x \text{ yang lain} \end{cases}$

2.7 Regresi Nonparametrik Kernel

Regresi kernel adalah teknik statistika nonparametrik untuk mengestimasi fungsi regresi m pada model regresi nonparametrik $y_i = m(t_i) + \varepsilon_i$. Misalkan diberikan data sebagai berikut: $\{t_i, y_i\}$, i=1,2,…,n, maka nilai m(t) ekuivalen dengan nilai harapan dari variabel respon apabila variabel $T = t$ telah diketahui. Diasumsikan variabel respon dan prediktor adalah variabel random. Secara matematis ditulis:

$$m(t) = E(Y|T = t) = \int_{-\infty}^{\infty} y f(y|t) dy = \int_{-\infty}^{\infty} y \frac{f(t,y)}{f(t)} dy \qquad (5)$$

Dengan mengadopsi estimator densitas kernel, m(t) dapat ditulis sebagai berikut :

$$m(t) = \frac{\int_{-\infty}^{\infty} y \hat{f}_{h_t,h_y}(t,y) dy}{\hat{f}_{h_t}(t)} \qquad (6)$$

Dengan $h_t = h$, didapatkan estimator Nadaraya-Watson sebagai berikut:

$$\hat{m}(t) = \frac{\sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right) y_i}{\sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right)} \qquad (7)$$

$$\hat{m}(t) = \sum_{i=1}^{n} W_i(t)\, y_i \qquad (8)$$

dengan

$$W_i(t) = \frac{K\left(\frac{t - t_i}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right)} \qquad (9)$$

Sehingga persamaan regresi nonparametrik diperoleh :

$$\hat{y}_i = \hat{m}\,(t_i) = \sum_{j=1}^{n} W_j(t_i)\, y_j$$

Jika ditulis dalam bentuk matriks, $\hat{y}_i$ dapat ditulis sebagai berikut :

$$\boldsymbol{\hat{Y}} = \boldsymbol{\hat{M}(t)} = \boldsymbol{W(t)\,Y}$$

2.8 Regresi Semiparametrik

Regresi semiparametrik dapat digunakan apabila dalam suatu kasus terdiri dari komponen parametrik yang diketahui polanya dan komponen nonparametrik yang *smooth* tidak diketahui polanya[6]. Bentuk umum model regresi semiparametrik dapat dituliskan sebagai berikut:

$$y_i = \sum_{k=1}^{p} \beta_k x_{ik} + m\,(t_i) + \varepsilon_i \quad , i = 1, 2, \ldots, n\ , k = 1, 2, \ldots, p \qquad (10)$$

dengan:

y_i : Variabel respon pengamatan ke-i

x_{ik} : Variabel bebas parametrik ke-k pengamatan ke-i

t_i : Variabel bebas nonparametrik pengamatan ke-i

β_k : Parameter variabel bebas x_k

m : komponen nonparametrik dengan fungsi regresi yang tidak diketahui
ε_i : residual ke-i
Persamaan (10) dapat juga ditulis menjadi $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{M(t)} + \boldsymbol{\varepsilon}$

2.9 Estimasi Model Semiparametrik Kernel

Pada model regresi semiparametrik pada persamaan (10), m dan β merupakan fungsi dan parameter yang akan diestimasi dari data. Estimasi m menggunakan estimator kernel dan estimasi parameter β menggunakan metode kuadrat terkecil. Misalkan jika terdapat n observasi dengan p variabel bebas x dan 1 variabel bebas t, maka persamaan (12) dapat ditulis sebagai berikut:

$$y_i - \sum_{k=1}^{p} \beta_k x_{ik} = m(t_i) + \varepsilon_i \tag{11}$$

Jika $y_i{}^* = y_i - \sum_{k=1}^{p} \beta_k x_{ik}$, maka nilai m(t) ekuivalen dengan nilai harapan dari variabel respon apabila variabel $T = t$ telah diketahui. Diasumsikan variabel respon dan prediktor adalah variabel random. Secara matematis ditulis:

$$m(t) = E(y^*|T = t) = \int_{-\infty}^{\infty} y^* f(y|t) dy = \frac{\int_{-\infty}^{\infty} y^* f(t,y) dy^*}{f(t)} \tag{12}$$

Selanjutnya dengan menggunakan estimator Nadaraya-Watson seperti regresi nonparametrik kernel diperoleh persamaan sebagai berikut:

$$\widehat{m}(t) = \frac{\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right) y_i{}^*}{\sum_{i=1}^{n} K\left(\frac{t-t_i}{h}\right)} = \sum_{i=1}^{n} W_i(t) \left(y_i - \sum_{k=1}^{p} \beta_k x_{ik}\right) \tag{13}$$

Sehingga diperoleh $\widehat{m}(t)$ untuk menduga $m(t)$ dengan model estimasi.

$$y_i = \sum_{k=1}^{p} \beta_k x_{ik} + \widehat{m}(t_i) + \varepsilon_i \tag{14}$$

Persamaan (14) juga dapat ditulis menjadi $\boldsymbol{Y} = \boldsymbol{X\beta} + \widehat{\boldsymbol{M}}(\boldsymbol{t}) + \varepsilon$. Dengan metode kuadrat terkecil diperoleh estimasi β dengan meminimumkan jumlah kuadrat *error*nya, yaitu sebagai berikut:

$$\begin{aligned}
\boldsymbol{\varepsilon^T\varepsilon} &= \left[\boldsymbol{Y} - \boldsymbol{X\beta} - \widehat{\boldsymbol{M}}(\boldsymbol{t})\right]^{\boldsymbol{T}} \left[\boldsymbol{Y} - \boldsymbol{X\beta} - \widehat{\boldsymbol{M}}(\boldsymbol{t})\right] \\
&= [\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{W(t)}(\boldsymbol{Y} - \boldsymbol{X\beta})]^{\boldsymbol{T}} [\boldsymbol{Y} - \boldsymbol{X\beta} - \boldsymbol{W(t)}(\boldsymbol{Y} - \boldsymbol{X\beta})] \\
&= [(\boldsymbol{Y} - \boldsymbol{W(t)Y}) - (\boldsymbol{X} - \boldsymbol{W(t)X})\boldsymbol{\beta}]^{\boldsymbol{T}} [(\boldsymbol{Y} - \boldsymbol{W(t)Y}) - (\boldsymbol{X} - \boldsymbol{W(t)X})\boldsymbol{\beta}]
\end{aligned}$$

Misalkan $\widetilde{\boldsymbol{Y}} = \boldsymbol{Y} - \boldsymbol{W(t)Y}$ dan $\widetilde{\boldsymbol{X}} = \boldsymbol{X} - \boldsymbol{W(t)X}$, maka:

$$\begin{aligned}
\boldsymbol{\varepsilon^T\varepsilon} &= (\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta})^{\boldsymbol{T}} (\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{X}}\boldsymbol{\beta}) \\
&= \widetilde{\boldsymbol{Y}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}} - \boldsymbol{2\beta^T}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}} + \boldsymbol{\beta^T}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}\boldsymbol{\beta}
\end{aligned}$$

$\frac{\partial \varepsilon^T \varepsilon}{\partial \beta} = 0$, maka diperoleh $-\boldsymbol{2}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}} + \boldsymbol{2}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}} = \boldsymbol{0}$

$\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}} = \widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}\widehat{\boldsymbol{\beta}}$

$\frac{\partial^2 \varepsilon^T \varepsilon}{\partial \beta \partial \beta^T} = \boldsymbol{2}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}$ merupakan matriks semidefinit positif. Jika $\boldsymbol{2}\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}}$ adalah matriks definit positif, maka $\hat{\beta}$ merupakan estimator kuadrat terkecil untuk β dengan:

$$\hat{\beta} = (\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}})^{-\boldsymbol{1}} \widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}} \tag{15}$$

Berdasarkan uraian di atas, jika $\hat{y}_i$ adalah variabel penduga untuk y_i diperoleh model dugaan semiparametrik kernel yaitu $\hat{y}_i = \sum_{k=1}^{p} \hat{\beta}_k x_{ik} + \widehat{m}(t_i)$ dengan
$\widehat{m}(t_i) = \sum_{j=1}^{n} W_j(t_i) \left(y_j - \sum_{k=1}^{p} \hat{\beta}_k x_{ik}\right)$ dan $\hat{\beta} = (\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}})^{-\boldsymbol{1}} \widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{Y}}$.
Sementara itu persamaan (10) juga dapat dinyatakan dengan

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{M}}(\boldsymbol{t}) = \left[\boldsymbol{W(t)} + \boldsymbol{P}_{\widetilde{\boldsymbol{x}}}\left(\boldsymbol{I} - \boldsymbol{W(t)}\right)\right]\boldsymbol{Y} = \boldsymbol{H(h)Y}$$

Dengan $\boldsymbol{P}_{\widetilde{\boldsymbol{x}}} = \widetilde{\boldsymbol{X}}(\widetilde{\boldsymbol{X}}^{\boldsymbol{T}}\widetilde{\boldsymbol{X}})^{-\boldsymbol{1}} \widetilde{\boldsymbol{X}}^{\boldsymbol{T}}$

2.10 Pemilihan *Bandwidth*

Pemilihan *bandwidth* optimal memiliki peranan penting dalam analisis regresi. Salah satu metode yang digunakan dalam menentukan *bandwidth* optimal, yaitu dengan *Generalized Cross Validation* (GCV)[15]. Jika Y diestimasi maka diperoleh:

$$\hat{\boldsymbol{Y}} = \boldsymbol{H}(\boldsymbol{h})\,\boldsymbol{Y}$$

Dimana $\boldsymbol{H}(\boldsymbol{h}) = \boldsymbol{W}(\boldsymbol{t}) + \boldsymbol{P}_{\tilde{\boldsymbol{x}}}(\boldsymbol{I} - \boldsymbol{W}(\boldsymbol{t}))$ dengan $\boldsymbol{P}_{\tilde{\boldsymbol{x}}} = \tilde{\boldsymbol{X}}(\tilde{\boldsymbol{X}}^{\boldsymbol{T}}\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}^{\boldsymbol{T}}$

$$GCV(h) = \frac{n^{-1}\sum_{i=1}^{n}[y_i - \hat{y}_i]^2}{\{1 - tr(\boldsymbol{H}(\boldsymbol{h}))/n\}^2}, i = 1,2,\dots,n$$

Dengan $GCV(h)$: nilai GCV pada *bandwidth* h
n : banyak subjek
y_i : data aktual subjek ke-i
$\hat{y}_i$: hasil estimasi subjek ke-i
$tr(H)$: jumlah dari elemen diagonal utama matriks penghalus ukuran $n\ x\ n$

Nilai *bandwidth* (h) yang optimal adalah nilai yang bersesuaian dengan GCV(h) yang minimum.

2.11 Evaluasi Ketepatan Model Regresi

a. Koefisien Determinasi

Adapun nilai dari koefisien determinasi dapat diperoleh dengan rumus:

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKG} \tag{16}$$

dengan : JKR = $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$
JKG = $\sum_{i=1}^{n}(y_i - \hat{y})^2$

dimana R^2: koefisien determinasi
y_i : data aktual subjek ke-i
$\hat{y}_i$: hasil estimasi subjek ke-i
$\bar{y}$: rata- rata dari data aktual

Nilai dari koefisien determinasi terletak antara 0 dan 1 ($0 \leq R^2 \leq 1$). Semakin mendekati 1 nilai dari R^2 maka semakin baik model yang dihasilkan dan semakin mendekati angka 0 maka semakin kurang baik model yang dihasilkan[14].

b. MAPE

Nilai tengah kesalahan persentase absolute (Mean Absolute Error) dapat dinyatakan sebagai berikut[9]:

$$MAPE = \sum_{i=1}^{n}\frac{\left|\frac{y_i - \hat{y}_i}{y_i}\right|}{n} \times (100\%)$$

Nilai MAPE yang dihasilkan mempunyai interpretasi sebagai berikut[3]:

a. MAPE < 10% : peramalan sangat akurat
b. 10% ≤ MAPE < 20% : peramalan tersebut baik
c. 20% ≤ MAPE <50% : peramalan masih dalam kewajaran
d. MAPE ≥ 50% : peramalan tidak akurat

3. METODE PENELITIAN

3.1 Sumber dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu IHSG, jumlah uang beredar, dan kurs rupiah terhadap dollar. Data yang digunakan meliputi data *in sample* dan *out sampel*. Dimana data *in sample* merupakan data bulanan mulai bulan Januari 2010 sampai bulan Agustus 2015, sedangkan data *out sampel* merupakan data bulanan mulai bulan September 2015 sampai bulan Februari 2016. Data IHSG tersebut dapat diakses pada situs www.yahoo.finance.com. Data jumlah uang beredar diperoleh dari Badan Pusat Statistik. Data kurs rupiah terhadap dollar diperoleh dari Bank Indonesia. Variabel yang yang digunakan pada penelitian ini terdiri dari variabel respon dan variabel prediktor, variabel-variabel tersebut sebagai berikut:

1. Indeks harga saham gabungan sebagai variabel respon (y)
2. Jumlah uang beredar sebagai variabel prediktor 1 (x)
3. Kurs rupiah terhadap dollar sebagai variabel prediktor 2 (t)

3.2 Langkah-langkah Analisis Data

Software statistik yang digunakan pada penelitian ini adalah R 3.0.3, Minitab, dan Microsoft Excel 2007. Berikut langkah-langkah yang dilakukan:

1. Menetapkan variabel
2. Mengumpulkan data yaitu data IHSG, jumlah uang beredar, dan kurs rupiah terhadap dollar.
3. Menentukan data *in sample* dan *out sample*.
4. Melakukan deskripsi pada data yang dikumpulkan dengan cara membuat *scatterplot* untuk masing-masing variabel prediktor dengan variabel respon.
5. Menentukan variabel komponen parametrik dan komponen nonparametrik.
6. Menentukan fungsi kernel yang digunakan, yaitu kernel gaussian dan kernel segitiga.
7. Menentukan besar *bandwidth*.
8. Estimasi parameter β.
9. Pemilihan model optimal dengan metode GCV.
10. Mengestimasi model semiparametrik.
11. Menentukan nilai MAPE untuk data *out sample* dan koefisien determinasi (R^2) untuk data *in sample*.
12. Menyimpulkan hasil penelitian.

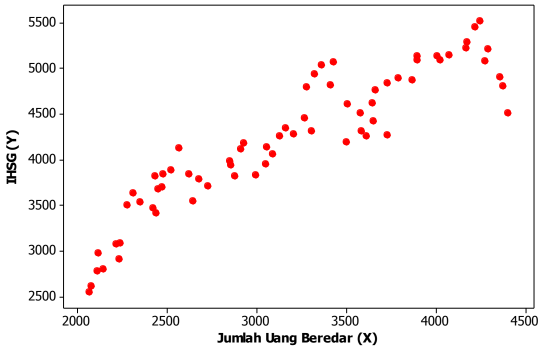
4. HASIL DAN PEMBAHASAN

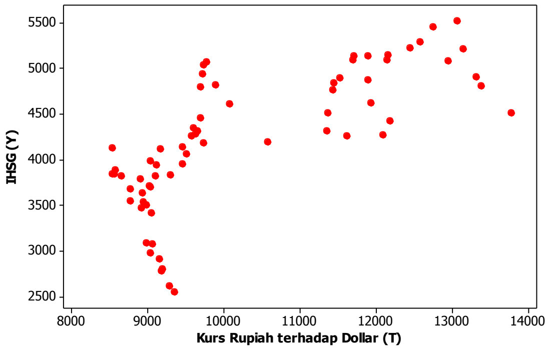
4.1 Pemodelan IHSG dengan Regresi Semiparametrik

Faktor-faktor yang mempengaruhi IHSG tidaklah selalu memiliki kurva regresi tertentu sehingga dalam pemodelannya menggunakan regresi semiparametrik kernel. Dimana di dalam regresi semiparametrik terdapat dua komponen, yaitu komponen parametrik dan komponen nonparametrik. Berikut ini adalah langkah-langkah yang dilakukan untuk mendapatkan model regresi semiparametrik :

4.1.1 Plot antara IHSG dengan Faktor-Faktor yang Mempengaruhinya

Berikut ini adalah gambar beberapa *scatterplot* antara variabel respon (Y) dengan masing-masing variabel prediktor (X dan T):





Berdasarkan plot data di atas, dapat dilihat bahwa plot data antara variabel respon IHSG (y) dengan variabel Jumlah Uang Beredar (x) membentuk pola sebaran titik yang linier, sedangkan plot data antara variabel IHSG dengan variabel kurs rupiah terhadap dollar (t) menyebar atau tidak berpola.

4.1.2 Model Regresi Semiparametrik Kernel

Model regresi semiparametrik kernel pada variabel x sebagai komponen parametrik dan variabel t sebagai komponen nonparametrik adalah sebagai berikut:

$$\hat{y}_i = \hat{\beta}x_i + \hat{m}\,(t_i)$$

4.1.3 Pemilihan *Bandwidth* Optimal

Untuk mendapatkan *bandwidth* yang optimal, digunakan metode *Generalized Cross Validation* (GCV). Pemilihan tersebut dilakukan dengan *trial and error* hingga diperoleh nilai GCV yang paling minimum.

4.1.3.1 Pemodelan Regresi Semiparametrik dengan Fungsi Kernel Gaussian

Dengan menggunakan kernel gaussian yang diolah menggunakan program R, dilakukan pengujian *bandwidth* secara *trial and error*. Berikut diberikan beberapa nilai GCV terkecil dari *bandwidth* yang dicobakan beserta nilai parameter β yang diperoleh.

No.	*Bandwidth* (h)	GCV	Parameter β
1	47,92	34675,27153	1,431653373
2	47,93	34675,27060	1,431653646
3	**47,94**	**34675,27047**	**1,431653920**
4	47,95	34675,27112	1,431654194
5	47,96	34675,27257	1,431654469

Dapat dilihat bahwa GCV terkecil terdapat pada *bandwidth* 47,94 dengan nilai GCV sebesar 34675,27047. Sehingga diperoleh *bandwidth* optimal dengan menggunakan fungsi kernel gaussian yaitu 47,94.

4.1.3.2 Pemodelan Regresi Semiparametrik dengan Fungsi Kernel Segitiga

Dengan menggunakan kernel segitiga yang diolah menggunakan program R, dilakukan pengujian *bandwidth* secara *trial and error*. Berikut diberikan beberapa nilai GCV terkecil dari *bandwidth* yang dicobakan beserta nilai parameter β yang diperoleh.

No.	*Bandwidth* (h)	GCV	Parameter β
1	216,71	36004,78284	1,426199045
2	216,72	36004,78270	1,426199832
3	**216,73**	**36004,78264**	**1,426200619**
4	216,74	36004,78265	1,426201406
5	216,75	36004,78275	1,426202193

Dapat dilihat bahwa GCV terkecil terdapat pada *bandwidth* 216,73 dengan nilai GCV sebesar 36004,78264. Sehingga diperoleh *bandwidth* optimal dengan menggunakan fungsi kernel segitiga yaitu 216,73.

4.1.4 Pemodelan Regresi Semiparametrik Kernel dengan *Bandwidth* Optimal

Bandwidth optimal yang diperoleh dari nilai GCV yang terkecil digunakan dalam pemodelan regresi semiparametrik kernel. Model regresi semiparametrik kernel adalah sebagai berikut:

$$\hat{y}_i = \hat{\beta}x_i + \hat{m}\,(t_i)$$

Persamaan model regresi semiparametrik kernel gaussian adalah sebagai berikut :

$$\hat{y}_i = 1{,}431653920\, x_i + \frac{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{47{,}94}\right)\left(y_j - 1{,}431653920\, x_j\right)}{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{47{,}94}\right)}$$

Persamaan model regresi semiparametrik kernel segitiga adalah sebagai berikut :

$$\hat{y}_i = 1{,}426200619\, x_i + \frac{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{216{,}73}\right)\left(y_j - 1{,}426200619\, x_j\right)}{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{216{,}73}\right)}$$

4.1.5 Evaluasi Ketepatan Model

4.1.5.1 Koefisien Determinasi

a. R^2 Fungsi Kernel Gaussian

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKG} = \frac{36070215{,}62}{36070215{,}62 + 807001{,}2956} = 0{,}9781165$$

Berdasarkan nilai perhitungan koefisien determinasi (R^2) tersebut, dapat dijelaskan bahwa variabel independen x dan t mempengaruhi variabel dependen y sebesar 97,81% dan sebesar 2,19% dipengaruhi oleh variabel yang lainnya.

b. R^2 Fungsi Kernel Segitiga

$$R^2 = \frac{JKR}{JKT} = \frac{JKR}{JKR + JKG} = \frac{35291395{,}67}{35291395{,}67 + 1237393{,}778} = 0{,}9661255$$

Berdasarkan nilai perhitungan koefisien determinasi (R^2) tersebut, dapat dijelaskan bahwa variabel independen x dan t mempengaruhi variabel dependen y sebesar 96,61% dan sebesar 3,39% dipengaruhi oleh variabel yang lainnya.

4.1.5.2 MAPE

a. MAPE Fungsi Kernel Gaussian

$$MAPE = \left(\sum_{i=1}^{n} \frac{\left|\frac{y_i - \hat{y}_i}{y_i}\right|}{n}\right)(100\%) = 4{,}03625\ \%$$

Nilai MAPE kurang dari angka 10% yang berarti model regresi semiparametrik yang dihasilkan sangat akurat.

b. MAPE Fungsi Kernel Segitiga

$$MAPE = \left(\sum_{i=1}^{n} \frac{\left|\frac{y_i - \hat{y}_i}{y_i}\right|}{n}\right)(100\%) = 11{,}08122\ \%$$

Nilai MAPE berada di antara 10% - 20% yang berarti model regresi semiparametrik yang dihasilkan baik.

4.1.6 Perbandingan Model Regresi Semiparametrik Kernel antara Kernel Gaussian dan Segitiga

Fungsi Kernel	*Bandwidth*	GCV	R^2	MAPE
Gaussian	47,94	34675,27047	0,9781165	4,03625 %
Segitiga	216,73	36004,78264	0,9661255	11,08122 %

R^2 kernel gaussian lebih besar daripada R^2 kernel segitiga dan MAPE kernel gaussian lebih kecil daripada MAPE kernel segitiga. Sehingga dapat disimpulkan bahwa model regresi semiparametrik kernel gaussian lebih baik dibandingkan model regresi semiparametrik kernel segitiga.

5. KESIMPULAN

Berdasarkan hasil dari pembahasan, dapat disimpulkan bahwa persamaan model regresi semiparametrik kernel yang diperoleh untuk menduga IHSG dengan fungsi kernel gaussian adalah sebagai berikut

$$\hat{y}_i = 1{,}431653920\ x_i + \frac{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{47{,}94}\right)\left(y_j - 1{,}431653920\ x_j\right)}{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{47{,}94}\right)}$$

Sedangkan, Persamaan model regresi semiparametrik kernel yang diperoleh untuk menduga IHSG dengan fungsi kernel segitiga adalah sebagai berikut

$$\hat{y}_i = 1{,}426200619\ x_i + \frac{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{216{,}73}\right)\left(y_j - 1{,}426200619\ x_j\right)}{\sum_{j=1}^{n} K\left(\frac{t_i - t_j}{216{,}73}\right)}$$

Serta, model regresi semiparametrik kernel gaussian lebih baik dibandingkan model regresi semiparametrik kernel segitiga.

DAFTAR PUSTAKA

[1] Arifin, I., Hadi, G.W. 2009. *Membuka Cakrawala Ekonomi*. Jakarta: PT. Setia Purna.

[2] Carmona, A.R. 2004. *Statistical Analysis of Financial Data in S-Plus*. United States of America: Springer.

[3] Chen, R.J.C., Bloomfield, P., dan Cubbage, F.W. 2007. *Comparing Forcasting Models in Tourism*. Journal of Hospitality & Tourism Research 2007. DOI: 10.1177/ 1096348007309566

[4] Gujarati, D. 2003. *Basic Econometrics*. United States of America: McGraw-Hill.

[5] Hardle, W. 1994. *Applied Nonparametric Regression*. Berlin: Humboldt University.

[6] Hardle, W., Liang. H., Gao. J. 2000. *Partially Linear Model*. Heidelberg: Physica-Verlag.

[7] Husnan, S. 2003. *Dasar-Dasar Portofolio dan Analisis Sekuritas*. Yogyakarta: UPP-AMP YKPN.

[8] Lawrence, S.S. 2013. *Pengaruh Variabel Makro Ekonomi dan Harga Komoditas Terhadap Indeks Harga Saham Gabungan di Indonesia*. Jurnal FINESTA Vol. 1, No.2: Hal. 18-23.

[9] Makridakis, S., Wheelewright, S.C., McGee, V.E. 1995. *Metode dan Aplikasi Peramalan*. Diterjemahkan oleh: Untung Sus Andriyanto dan Abdul Basith. Jakarta: Erlangga. Terjemahan dari: Forecasting Methods and Applications.

[10] Ogden, R.T. 1997. *Essential Wavelets for Statistical Applications and Data Analysis*. Boston: Birkhauser.

[11] Samsul, M. 2006. *Pasar Modal & Manajemen Portofolio*. Jakarta: Erlangga.

[12] Sukirno, S. 2002. *Pengantar Teori Makroekonomi Edisi Kedua*. Jakarta: PT Raja Grafindo Persada

[13] Sunariyah. 2003. *Pengantar Pengetahuan Pasar Modal*. Jakarta: Erlangga.

[14] Widarjono, A. 2010. *Analisis Statistika Multivariat Terapan*. Yogyakarta: Unit Penerbit dan Percetakan STIM YKPN.

[15] Wu, H., Zhang, J.T. 2006. *Nonparametric Regression Methods for Longitudinal Data Analysis*. New York: John Wiley and Sons, Inc.